

A VARIATIONAL NONLINEAR HAUSDORFF–YOUNG INEQUALITY IN THE DISCRETE SETTING

DIOGO OLIVEIRA E SILVA

ABSTRACT. Following the works of Lyons [9, 10] and Oberlin, Seeger, Tao, Thiele and Wright [14], we relate the variation of certain discrete curves on the Lie group $SU(1, 1)$ to the corresponding variation of their linearized versions on the Lie algebra. Combining this with a discrete variational Menshov–Paley–Zygmund theorem, we establish a variational Hausdorff–Young inequality for a discrete version of the nonlinear Fourier transform on $SU(1, 1)$.

1. INTRODUCTION

The authors of [14] considered a continuous version of the $SU(1, 1)$ -valued nonlinear Fourier transform, and established a variational Hausdorff–Young inequality which served as inspiration for much of the present work. In this paper, we investigate the corresponding discrete problem and establish Theorem 1.1 below. Before stating it precisely, we introduce some context and background material.

1.1. The generalized special unitary group and its Lie algebra. Following Tao and Thiele [16], we are interested in a particular nonlinear Fourier transform taking values in the generalized special unitary group

$$SU(1, 1) = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C} \text{ and } |a|^2 - |b|^2 = 1 \right\}.$$

This is a connected three-dimensional Lie group which is isomorphic to $SL(2, \mathbb{R})$, and therefore neither compact nor simply connected. The corresponding Lie algebra is given by

$$\mathfrak{su}(1, 1) = \left\{ \begin{pmatrix} ir & s - it \\ s + it & -ir \end{pmatrix} : r, s, t \in \mathbb{R} \right\}.$$

Sometimes we shall identify a matrix of the form $\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}$ with its first row, thereby writing expressions like $(a, b) \in SU(1, 1)$. The operator norm of a matrix $(a, b) \in SU(1, 1)$ acting on the Hilbert space \mathbb{C}^2 is given by the simple expression

$$\|(a, b)\|_{\text{op}} = |a| + |b|, \tag{1.1}$$

see [16, Lemma 33]. We endow $SU(1, 1)$ with the distance

$$d(X, Y) = \log(1 + \|X^{-1} \cdot Y - I\|_{\text{op}}), \tag{1.2}$$

Date: April 4, 2017.

2010 Mathematics Subject Classification. 43A32, 43A50.

Key words and phrases. Nonlinear Fourier transform, variational estimates, Hausdorff–Young inequality. The author was partially supported by the Hausdorff Center for Mathematics and DFG grant CRC 1060.

where I denotes the 2×2 identity matrix. We also turn $\mathfrak{su}(1, 1)$ into a normed algebra by endowing it with the operator norm. For a further discussion of the Lie group $SU(1, 1)$ and its Lie algebra, we refer the reader to [15, §10.4] and [16, Appendix 2].

1.2. The nonlinear Fourier transform. Let \mathcal{D} denote the open unit disc in the complex plane, whose boundary coincides with the unit circle \mathbb{T} . Given a square-integrable, \mathcal{D} -valued sequence $F \in \ell^2(\mathbb{Z}, \mathcal{D})$, its linear (inverse) Fourier transform is given by

$$\check{F}(z) = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} F_n z^n, \quad (z \in \mathbb{T})$$

where the limit is taken in the sense of $L^2(\mathbb{T})$. With the goal of defining a nonlinear analogue of the Fourier transform, consider a compactly supported, \mathcal{D} -valued sequence F , which will often be referred to as a *potential*. The associated transfer matrices $\{T_n\}$ are defined, for each $n \in \mathbb{Z}$ and $z \in \mathbb{T}$, by

$$T_n(z) = \begin{pmatrix} 1 & F_n z^n \\ F_n z^{-n} & 1 \end{pmatrix} (1 - |F_n|^2)^{-\frac{1}{2}}.$$

Note that $T_n(z) \in SU(1, 1)$. We define the nonlinear Fourier transform of the potential F as an $SU(1, 1)$ -valued function on the unit circle \mathbb{T} , given by the expression

$$(a, b)(z) = \lim_{N \rightarrow \infty} \prod_{n=-N}^N T_n(z), \quad (z \in \mathbb{T}) \quad (1.3)$$

where the ordered product is actually finite. In this case, a contour integration establishes the following nonlinear analogue of Plancherel's $L^2 - L^2$ estimate:¹

$$\int_{\mathbb{T}} \log |a(z)| = -\frac{1}{2} \sum_{n \in \mathbb{Z}} \log(1 - |F_n|^2), \quad (1.4)$$

which can then be used to extend the definition of (a, b) to \mathcal{D} -valued sequences F which are merely assumed to be square-integrable. If the sequence F is further required to belong to ℓ^1 , then a clever but elementary convexity argument establishes the estimate

$$\sup_{z \in \mathbb{T}} \sqrt{\log |a(z)|} \leq \sum_{n \in \mathbb{Z}} \sqrt{\log((1 - |F_n|^2)^{-\frac{1}{2}})}, \quad (1.5)$$

which should be thought of as a nonlinear equivalent of the trivial $L^1 - L^\infty$ estimate for the linear Fourier transform. One would like to interpolate estimates (1.4) and (1.5) to conclude a nonlinear version of the $L^p - L^{p'}$ Hausdorff–Young inequality when $1 < p < 2$. Here and throughout the paper, $p' = \frac{p}{p-1}$ denotes the exponent conjugate to p . Even though standard interpolation tools are not available in the present nonlinear context, the following estimate is a consequence of the seminal work of Christ and Kiselev [1, 2] on the spectral theory of one-dimensional Schrödinger operators: Given $1 < p < 2$, there exists a constant $C_p < \infty$ such that, for every sequence $F \in \ell^p(\mathbb{Z}, \mathcal{D})$,

$$\left\| \sqrt{\log |a(z)|} \right\|_{L^{p'}(\mathbb{T})} \leq C_p \left\| \sqrt{|\log(1 - |F_n|^2)|} \right\|_{\ell^p(\mathbb{Z})}. \quad (1.6)$$

¹Here and in what follows, $\int_{\mathbb{T}}$ denotes the integral with respect to Lebesgue measure on \mathbb{T} , normalized to have total mass 1.

The cases $p = 1$ and $p = 2$ of inequality (1.6) boil down to (1.5) and (1.4), respectively. For precise statements and proofs of these estimates, see [16] and the references therein. An interesting open problem is whether the constant C_p in inequality (1.6) can be chosen uniformly in p as $p \rightarrow 2^-$. The answer is known to be affirmative in a variant for a particular Cantor group model of the continuous nonlinear Fourier transform [6], see also [7] for a recent related result.

1.3. Main result. In this paper, we aim at a variational refinement of inequality (1.6). With this purpose in mind, consider the following truncated version of the linear and the nonlinear Fourier transforms of the potential F , respectively denoted by $\sigma(F)$ and $\gamma[F]$, given at height $N \in \mathbb{Z}$ by

$$\sigma(F)(N; z) = \sum_{n=-\infty}^N \begin{pmatrix} 0 & F_n z^n \\ \overline{F_n} z^{-n} & 0 \end{pmatrix}, \quad \text{and} \quad \gamma[F](N; z) = \prod_{n=-\infty}^N T_n(z). \quad (1.7)$$

For each fixed $z \in \mathbb{T}$, it will be convenient to think of the maps $N \mapsto \gamma[F](N; z)$ and $N \mapsto \sigma(F)(N; z)$ as discrete curves taking values on $\mathrm{SU}(1, 1)$ and $\mathfrak{su}(1, 1)$, respectively. To make the notation less cumbersome, we shall write $\gamma_N(z)$ instead of $\gamma[F](N; z)$ whenever there is no danger of confusion, and similarly for the curve $\sigma(F)$. Given an exponent $1 \leq r \leq \infty$, we will be interested in measuring the r -variation in the variable N of the curves $\sigma = \sigma(F)$ and $\gamma = \gamma[F]$. If $1 \leq r < \infty$, then these variations at a given point $z \in \mathbb{T}$ are defined as

$$\begin{aligned} \mathcal{V}_r(\sigma)(z) &= \sup_K \sup_{N_0 < \dots < N_K} \left(\sum_{j=0}^{K-1} \|\sigma_{N_{j+1}}(z) - \sigma_{N_j}(z)\|_{\mathrm{op}}^r \right)^{\frac{1}{r}}, \\ \mathcal{V}_r(\gamma)(z) &= \sup_K \sup_{N_0 < \dots < N_K} \left(\sum_{j=0}^{K-1} d(\gamma_{N_j}(z), \gamma_{N_{j+1}}(z))^r \right)^{\frac{1}{r}}, \end{aligned}$$

where the suprema are taken over all strictly increasing finite sequences of integers $N_0 < N_1 < \dots < N_K$ and over all integers K . We can extend this to the case $r = \infty$ in the usual manner. The main result of this paper is the following theorem, which should be compared to the estimates in [14, p. 461].

Theorem 1.1. *Let $1 \leq p < 2$ and $r > p$. Then there exists a constant $C_{p,r} < \infty$ such that*

$$\|\mathcal{V}_r(\gamma[F])\|_{L^{p'}(\mathbb{S})} + \|\mathcal{V}_r(\gamma[F])\|_{L^{\frac{p'}{r}}(\mathbb{T} \setminus \mathbb{S})}^{\frac{1}{r}} \leq C_{p,r} \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^p(\mathbb{Z})}, \quad (1.8)$$

for every $F \in \ell^p(\mathbb{Z}, \mathcal{D})$. Here \mathbb{S} denotes the subset of the unit circle defined as follows:

$$\mathbb{S} = \mathbb{S}_{p,r}[F] = \{z \in \mathbb{T} : \mathcal{V}_s(\gamma[F])(z) \leq 1\}, \quad (1.9)$$

where $s = r$ if $p < r < 2$, and $s = \frac{p+2}{2}$ if $r \geq 2$.

A few remarks may help to further orient the reader.

- The usual logarithm (that appears for instance on the left-hand side of inequality (1.6)) is hidden in the metric d which we have placed on the group $\mathrm{SU}(1, 1)$.

- The condition $r > p$ is sharp, except at the endpoint $p = 1$. In this case, if $p = r = 1$, then inequality (1.8) holds with $C_{1,1} \leq 1$. In the general case, (1.8) can only hold if $r \geq p$, as can easily be seen by considering potentials given by $F_n = \frac{1}{2}$ if $|n| \leq N$, and $F_n = 0$ otherwise. Furthermore, if $p > 1$, then (1.8) can only hold under the strict inequality $r > p$, as discussed after the statement of the corresponding linearized result, Proposition 3.1 below.
- Extending Theorem 1.1 to $p = 2$, even when $r = \infty$, is a challenging open problem that would imply a nonlinear analogue of Carleson's theorem on almost everywhere convergence of partial Fourier series, see [12, 13] for an extended discussion. More modestly, one can ask for a restricted weak type result at the endpoint $r = p$ when $1 < p < 2$.

Let us briefly comment on the proof of Theorem 1.1. It comprises two parts which constitute the upcoming sections. In §2, we follow the adaptation of the work of Lyons [9, 10] by Oberlin et al. [14] to study variation norms on $SU(1, 1)$. In particular, given a potential F , we show that the r -variation of the discrete curve $\gamma[F]$ can be controlled by the r -variation of the linearized curve $\sigma(F)$ plus an extra term that accounts for the possible presence of large jumps, as long as $r < 2$. Most of the analysis is local, and uses a modified induction on scale (or “Bellman function”) argument inspired by [14, Lemma C.3] to control multiple steps of the curve $\gamma[F]$ at once. The corresponding step from [14] relies on a key subdivision result, [14, Lemma C.2], which asserts that any continuous curve can be decomposed into two proper sub-curves, each of which satisfies good r -variational bounds. No such decomposition is available in the discrete setting. Instead, we decompose the partial sum process into a part encoding large range displacement and a part which tracks variation on a prescribed local scale, and analyze them separately. This strategy already appeared in the works [8, 11]. One is then left with establishing a discrete variational version of the classical Menshov–Paley–Zygmund theorem. This step requires $r > p$, and is accomplished in §3 via an adaptation of the original argument of Christ–Kiselev [2] to the variational setting. The proof of Theorem 1.1 is then completed in a few strokes in §4.

Notation. If x, y are real numbers, we write $x = O(y)$ or $x \lesssim y$ if there exists a finite absolute constant C such that $|x| \leq C|y|$. If we want to make explicit the dependence of the constant C on some parameter α , we write $x = O_\alpha(y)$ or $x \lesssim_\alpha y$. We also write $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$. The indicator function of a set E is denoted by $\mathbf{1}_E$.

2. VARIATIONAL ESTIMATES

This section is devoted to controlling the variation of a curve on the Lie group $SU(1, 1)$ in terms of the variation of its linearized version. We refer the reader to [3, 4, 5, 14, 17] for background on variation seminorms, and proceed to explore some elementary properties which will be useful in our analysis. We shall state them for the exponentiated curve $\gamma = \gamma[F]$ only, but in each case it will be clear what the corresponding property for the linearized curve $\sigma = \sigma[F]$ should be. To make the forthcoming notation less cumbersome, we will occasionally drop the dependence on F and z . An application of Minkowski's inequality reveals that the r -variation is decreasing in the exponent r :

$$\mathcal{V}_s(\gamma) \leq \mathcal{V}_r(\gamma), \text{ whenever } 1 \leq r \leq s \leq \infty. \quad (2.1)$$

Implicit in the notation $\mathcal{V}_r(\gamma)$ is the fact that the variation is taken over the whole domain of definition of γ . We will sometimes need to restrict the domain of the discrete curves under consideration. Given integers $a \leq b$, the restriction of the curve γ to an interval $[a, b] \subset \mathbb{Z}$ will be denoted by $\gamma|_{[a, b]}$, and the r -variation of the curve γ on the interval $[a, b]$ is defined as

$$\mathcal{V}_r(\gamma; [a, b]) = \sup_K \sup_{a=N_0 < \dots < N_K=b} \left(\sum_{j=0}^{K-1} d(\gamma_{N_j}, \gamma_{N_{j+1}})^r \right)^{\frac{1}{r}}.$$

Two discrete curves $\gamma_1 : [a, c] \rightarrow \text{SU}(1, 1)$ and $\gamma_2 : [c, b] \rightarrow \text{SU}(1, 1)$ can be concatenated provided $\gamma_1(c) = \gamma_2(c)$. If $1 \leq r \leq \infty$ and $\gamma_1 \oplus \gamma_2 : [a, b] \rightarrow \text{SU}(1, 1)$ denotes the concatenation of γ_1 and γ_2 , then it is a straightforward matter to establish the triangle inequalities

$$(\mathcal{V}_r^r(\gamma_1) + \mathcal{V}_r^r(\gamma_2))^{\frac{1}{r}} \leq \mathcal{V}_r(\gamma_1 \oplus \gamma_2) \leq \mathcal{V}_r(\gamma_1) + \mathcal{V}_r(\gamma_2). \quad (2.2)$$

If $r = 1$, then these inequalities combine into an equality. Moreover, both inequalities in (2.2) extend to an arbitrary (even countably infinite) number of summands. We seek to establish the following quantitative result.

Proposition 2.1. *Let $1 \leq r < 2$. Then there exists a constant $C_r < \infty$ such that*

$$\mathcal{V}_r(\gamma)(z) \leq \mathcal{V}_r(\sigma)(z) + C_r \left(\mathcal{V}_r^2(\sigma)(z) \wedge \mathcal{V}_r^r(\sigma)(z) + \|F\|_{\ell^r(\mathbb{Z})}^{r-1} \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^r(\mathbb{Z})} \right), \quad (2.3)$$

for every potential F and every $z \in \mathbb{T}$.

Proof of Proposition 2.1. We lose no generality in assuming that $r > 1$. Indeed, the 1-variation \mathcal{V}_1 is additive as quantified by (2.2). Moreover, identity (1.1) and some straightforward algebra imply that the operator norm of the matrix $T_n(z)$ satisfies

$$\log(1 + \|T_n(z) - I\|_{\text{op}}) = \frac{1}{2} \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right). \quad (2.4)$$

These two observations imply

$$\mathcal{V}_1(\gamma)(z) = \sum_{n \in \mathbb{Z}} \mathcal{V}_1(\gamma; [n-1, n])(z) = \frac{1}{2} \sum_{n \in \mathbb{Z}} \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right), \text{ for every } z \in \mathbb{T},$$

which in turn implies a stronger form of (2.3). For $1 < r < 2$, the somewhat lengthy proof of Proposition 2.1 is divided into three steps. The first two steps deal with the local analysis.

Step 1. For every $1 < r < 2$, there exist constants $0 < \delta < \frac{1}{2}$ and $K < \infty$ with the following property. For every potential F , the curves $\gamma = \gamma[F]$ and $\sigma = \sigma(F)$ satisfy

$$\| \log(\gamma_M^{-1}(z) \cdot \gamma_N(z)) - (\sigma_N(z) - \sigma_M(z)) \|_{\text{op}} \leq K \mathcal{V}_r^2(\sigma; [M, N])(z) \quad (2.5)$$

for every $z \in \mathbb{T}$ and integers $M \leq N$, provided $\mathcal{V}_r(\sigma; [M, N])(z) \leq \delta$.

Start by noting that, if δ is sufficiently small, then the matrix $(\gamma_M^{-1} \cdot \gamma_N)(z) = \prod_{M < n \leq N} T_n(z)$ is sufficiently close to the identity so that the logarithm is well-defined. If $N = M + 1$, then the left-hand side of inequality (2.5) can be determined exactly. Indeed, a straightforward computation yields

$$\log(\gamma_{N-1}^{-1} \cdot \gamma_N) = g(F_N)(0, F_N z^N),$$

where the function g is given by $g(t) = (2|t|)^{-1} \log \left(\frac{1+|t|}{1-|t|} \right)$. It follows that

$$\|\log(\gamma_{N-1}^{-1} \cdot \gamma_N) - (\sigma_N - \sigma_{N-1})\|_{\text{op}} = |g(F_N) - 1| \|(0, F_N z^N)\|_{\text{op}} = |g(F_N) - 1| |F_N|. \quad (2.6)$$

A Taylor expansion of the function g to order 2 reveals that this term is $O(|F_N|^3)$ if $|F_N| \leq \frac{1}{2}$. To handle the case $N > M + 1$, we perform an induction on scale argument which follows the corresponding part of the proof of [14, Lemma C.3], with some differences that we highlight below. Let us first establish inequality (2.5) in the case when the quantity $\mathcal{V}_r(\sigma; [M, N])$ is sufficiently small, depending on the potential F . We have that

$$\begin{aligned} \log(\gamma_M^{-1} \cdot \gamma_N) + \frac{1}{2} \left(\sum_{M < n \leq N} \log(1 - |F_n|^2) \right) I &= \log \left(\prod_{M < n \leq N} \left(I + (0, F_n z^n) \right) \right) \\ &= \log(I + (\sigma_N - \sigma_M) + H_M^N), \end{aligned} \quad (2.7)$$

where $H_M^N = H_M^N[F]$ denotes the higher order terms that appear when we expand the product in the argument of the logarithm. Taylor expanding the logarithm, we see that

$$\|\log(I + (\sigma_N - \sigma_M) + H_M^N) - (\sigma_N - \sigma_M)\|_{\text{op}} = O(\|\sigma_N - \sigma_M\|_{\text{op}}^2), \quad (2.8)$$

provided $\|\sigma_N - \sigma_M\|_{\text{op}}$ is sufficiently small. On the other hand,

$$\|\sigma_N - \sigma_M\|_{\text{op}} \leq \mathcal{V}_r(\sigma; [M, N]). \quad (2.9)$$

Further notice that, if $|F_n| \leq \frac{1}{2}$ for every n , then

$$\frac{1}{2} \sum_{M < n \leq N} |\log(1 - |F_n|^2)| \leq \sum_{M < n \leq N} |F_n|^2 \leq \left(\sum_{M < n \leq N} |F_n|^r \right)^{\frac{2}{r}} \leq \mathcal{V}_r^2(\sigma; [M, N]). \quad (2.10)$$

Recalling (2.7), estimates (2.8), (2.9) and (2.10) imply that inequality (2.5) holds provided the quantity $\mathcal{V}_r(\sigma; [M, N])$ is sufficiently small, depending on the potential F . This provides the base of the induction scheme, which we now perform in order to remove the undesirable dependence on the potential. To specify the induction step, fix a constant of proportionality θ with $\frac{1}{2} < \theta < \frac{3}{4}$.

Inductive Claim. *If estimate (2.5) holds when $\mathcal{V}_r(\sigma; [M, N]) < \varepsilon$, for some $0 < \varepsilon < \delta$, then it also holds when $\mathcal{V}_r(\sigma; [M, N]) < \theta^{-\frac{1}{r}} \varepsilon$, provided K is chosen sufficiently large and δ is sufficiently small.*

We emphasize that both parameters K, δ are allowed to depend only on the exponent r , but that δ will be chosen as a function of K in the course of the proof. Let us proceed to prove the Inductive Claim. Consider a potential F and the corresponding curve $\sigma = \sigma(F)$, such that $A := \mathcal{V}_r(\sigma; [M, N])$ satisfies $A < \theta^{-\frac{1}{r}} \varepsilon$, for some $0 < \varepsilon < \delta$. In this case, we have in particular that $A < 2\delta$. By a *good subdivision* of the curve $\sigma|_{[M, N]}$ we mean a decomposition $\sigma|_{[M, N]} = \sigma_1 \oplus \sigma_2$ into discrete curves $\sigma_1 = \sigma|_{[M, L]}$ and $\sigma_2 = \sigma|_{[L, N]}$, such that the corresponding variations satisfy

$$\mathcal{V}_r(\sigma_1) \vee \mathcal{V}_r(\sigma_2) = \alpha^{\frac{1}{r}} A, \quad (2.11)$$

for some $1 - \theta \leq \alpha \leq \theta$. We split the analysis into two cases, according to whether or not a good subdivision of $\sigma|_{[M, N]}$ exists.

Suppose first that there exists a good subdivision $\sigma|_{[M,N]} = \sigma_1 \oplus \sigma_2$. In this case, the argument is analogous to the one in [14, Lemma C.3], but we repeat it here for the convenience of the reader. We can assume that the maximum in (2.11) is attained by the first piece, $\mathcal{V}_r(\sigma_1) = \alpha^{\frac{1}{r}} A$, from which it follows by the reverse triangle inequality that

$$\mathcal{V}_r(\sigma_2) \leq (1 - \alpha)^{\frac{1}{r}} A.$$

Since $\alpha \leq \theta$, we have that $\mathcal{V}_r(\sigma_1) < \varepsilon$, and therefore $\mathcal{V}_r(\sigma_2) < \varepsilon$ as well. By the induction hypothesis, it follows that

$$\|\log(\gamma_M^{-1} \cdot \gamma_L) - (\sigma_L - \sigma_M)\|_{\text{op}} \leq K(\alpha^{\frac{1}{r}} A)^2, \text{ and} \quad (2.12)$$

$$\|\log(\gamma_L^{-1} \cdot \gamma_N) - (\sigma_N - \sigma_L)\|_{\text{op}} \leq K((1 - \alpha)^{\frac{1}{r}} A)^2. \quad (2.13)$$

Estimate (2.12) and $A < 2\delta$ together imply

$$\begin{aligned} \|\log(\gamma_M^{-1} \cdot \gamma_L)\|_{\text{op}} &\leq \|\sigma_L - \sigma_M\|_{\text{op}} + K(\alpha^{\frac{1}{r}} A)^2 \leq \mathcal{V}_r(\sigma; [M, N]) + KA^2 \\ &= A(1 + KA) \leq A(1 + 2K\delta) = O(A), \end{aligned}$$

provided δ is chosen sufficiently small depending on K . Estimate (2.13) similarly implies

$$\|\log(\gamma_L^{-1} \cdot \gamma_N)\|_{\text{op}} = O(A).$$

Invoking the simple estimates

$$\|\exp(X) - I - X\|_{\text{op}} = O(\|X\|_{\text{op}}^2) \text{ and } \|\log(I + X) - X\|_{\text{op}} = O(\|X\|_{\text{op}}^2), \quad (2.14)$$

both of which hold with implicit constants less than 2 as long as the matrix X satisfies $\|X\|_{\text{op}} \leq \frac{1}{2}$, we then have that

$$\|\log(\gamma_M^{-1} \cdot \gamma_N) - \log(\gamma_M^{-1} \cdot \gamma_L) - \log(\gamma_L^{-1} \cdot \gamma_N)\|_{\text{op}} = O(A^2), \quad (2.15)$$

if δ is sufficiently small. Using the triangle inequality and estimates (2.12), (2.13) and (2.15), it follows that

$$\|\log(\gamma_M^{-1} \cdot \gamma_N) - (\sigma_N - \sigma_M)\|_{\text{op}} \leq K(\alpha^{\frac{2}{r}} + (1 - \alpha)^{\frac{2}{r}})A^2 + O(A^2). \quad (2.16)$$

Since $1 < r < 2$, we have that the quantity $\alpha^{\frac{2}{r}} + (1 - \alpha)^{\frac{2}{r}} < 1$ is bounded away from 1, for every $1 - \theta \leq \alpha \leq \theta$. If K is large enough, depending only on r , we can bound the right-hand side of (2.16) by KA^2 , and thus have established (2.5) if $A < \theta^{-\frac{1}{r}} \varepsilon$. This finishes the analysis of the case when a good subdivision exists.

Let us now assume that no good subdivision of the curve $\sigma|_{[M,N]}$ exists. In this case, there exists a one step sub-curve of $\sigma|_{[M,N]}$, say $\sigma|_{[L-1,L]}$, which accounts for most of its variation. Quantitatively,

$$\theta^{\frac{1}{r}} - (1 - \theta)^{\frac{1}{r}} < \frac{\mathcal{V}_r(\sigma; [L-1, L])}{\mathcal{V}_r(\sigma; [M, N])} = \frac{|F_L|}{A} \leq 1, \quad (2.17)$$

$$\mathcal{V}_r(\sigma; [M, L-1]) < (1 - \theta)^{\frac{1}{r}} A, \text{ and } \mathcal{V}_r(\sigma; [L, N]) < (1 - \theta)^{\frac{1}{r}} A. \quad (2.18)$$

Recalling (2.6), we have that

$$\|\log(\gamma_{L-1}^{-1} \cdot \gamma_L) - (\sigma_L - \sigma_{L-1})\|_{\text{op}} = |g(F_L) - 1||F_L|, \quad (2.19)$$

and therefore

$$\|\log(\gamma_{L-1}^{-1} \cdot \gamma_L)\|_{\text{op}} \leq |F_L| + O(|F_L|^3). \quad (2.20)$$

The first inequality in (2.18) implies $\mathcal{V}_r(\sigma; [M, L-1]) < \varepsilon$, and so by induction hypothesis

$$\|\log(\gamma_M^{-1} \cdot \gamma_{L-1}) - (\sigma_{L-1} - \sigma_M)\|_{\text{op}} \leq K((1-\theta)^{\frac{1}{r}} A)^2. \quad (2.21)$$

Since $A < 2\delta$, it follows that

$$\begin{aligned} \|\log(\gamma_M^{-1} \cdot \gamma_{L-1})\|_{\text{op}} &\leq \mathcal{V}_r(\sigma; [M, L-1]) + K(1-\theta)^{\frac{2}{r}} A^2 \\ &\leq (1-\theta)^{\frac{1}{r}} A(1 + 2K(1-\theta)^{\frac{1}{r}} \delta) \leq 2(1-\theta)^{\frac{1}{r}} A \leq \frac{2(1-\theta)^{\frac{1}{r}}}{\theta^{\frac{1}{r}} - (1-\theta)^{\frac{1}{r}}} |F_L|, \end{aligned} \quad (2.22)$$

provided δ is chosen sufficiently small as a function of K . An identical analysis applies to the sub-curve $\sigma|_{[L, N]}$ and yields

$$\|\log(\gamma_L^{-1} \cdot \gamma_N) - (\sigma_N - \sigma_L)\|_{\text{op}} \leq K((1-\theta)^{\frac{1}{r}} A)^2, \quad (2.23)$$

$$\|\log(\gamma_L^{-1} \cdot \gamma_N)\|_{\text{op}} \leq \frac{2(1-\theta)^{\frac{1}{r}}}{\theta^{\frac{1}{r}} - (1-\theta)^{\frac{1}{r}}} |F_L|. \quad (2.24)$$

From estimates (2.20), (2.22) and (2.24) it follows, as before, that

$$\|\log(\gamma_M^{-1} \cdot \gamma_N) - \log(\gamma_M^{-1} \cdot \gamma_{L-1}) - \log(\gamma_{L-1}^{-1} \cdot \gamma_L) - \log(\gamma_L^{-1} \cdot \gamma_N)\|_{\text{op}} = O(|F_L|^2) = O(A^2). \quad (2.25)$$

By the triangle inequality and estimates (2.19), (2.21), (2.23) and (2.25), we then have

$$\|\log(\gamma_M^{-1} \cdot \gamma_N) - (\sigma_N - \sigma_M)\|_{\text{op}} \leq |g(F_L) - 1| |F_L| + 2K(1-\theta)^{\frac{2}{r}} A^2 + O(A^2).$$

Since $2(1-\theta)^{\frac{2}{r}} < 1$, the right-hand side of this inequality can again be majorized by KA^2 , provided K is sufficiently large (depending on r). This concludes the analysis when no good subdivision exists. The Inductive Claim is now proved, and this settles Step 1.

Step 2. Given $1 < r < 2$, let δ be the constant promised by Step 1. Then, for every potential F , the curves $\gamma = \gamma[F]$ and $\sigma = \sigma(F)$ satisfy

$$\mathcal{V}_r(\gamma)(z) = \mathcal{V}_r(\sigma)(z) + O_r(\mathcal{V}_r^2(\sigma)(z)), \quad (2.26)$$

for every $z \in \mathbb{T}$, provided $\mathcal{V}_r(\sigma)(z) \leq \delta$.

The elementary estimates in (2.14) imply that short distances in the Lie group and the Lie algebra agree modulo a quadratically small error: Given a matrix $X \in \mathfrak{su}(1, 1)$ of sufficiently small operator norm,

$$d(I, \exp(X)) = \log(1 + \|\exp(X) - I\|_{\text{op}}) = \|X\|_{\text{op}} + O(\|X\|_{\text{op}}^2).$$

Consequently, if δ is small enough, then, for any integers $M \leq N$,

$$d(\gamma_M, \gamma_N) = d(I, \exp(\log(\gamma_M^{-1} \cdot \gamma_N))) = \|\log(\gamma_M^{-1} \cdot \gamma_N)\|_{\text{op}} + O(\|\log(\gamma_M^{-1} \cdot \gamma_N)\|_{\text{op}}^2).$$

In view of Step 1, we then have

$$d(\gamma_M, \gamma_N) = \|\sigma_N - \sigma_M\|_{\text{op}} + O(\mathcal{V}_r^2(\sigma; [M, N])).$$

Given $\varepsilon > 0$, choose an integer R which is large enough, so that $\mathcal{V}_r(\gamma) \leq \mathcal{V}_r(\gamma; [-R, R]) + \varepsilon$. For any partition $-R = N_0 < N_1 < \dots < N_K = R$, we may crudely estimate

$$O(\mathcal{V}_r^2(\sigma; [N_j, N_{j+1}])) \leq \mathcal{V}_r(\sigma; [-R, R]) O(\mathcal{V}_r(\sigma; [N_j, N_{j+1}])),$$

and conclude, by taking the ℓ^r sum in the j index, that

$$\left(\sum_{j=0}^{K-1} d(\gamma_{N_{j+1}}, \gamma_{N_j})^r \right)^{\frac{1}{r}} = \left(\sum_{j=0}^{K-1} \|\sigma_{N_{j+1}} - \sigma_{N_j}\|_{\text{op}}^r \right)^{\frac{1}{r}} + O(\mathcal{V}_r^2(\sigma; [-R, R])).$$

Taking the suprema over all partitions and over K , and then letting $\varepsilon \rightarrow 0^+$, we obtain the result. This concludes the proof of Step 2, and with it the local analysis is finished. The next and final step deals with the global part of the analysis.

Step 3. Global analysis and conclusion of the proof.

Given $1 < r < 2$, let δ be the constant promised by Step 1. Consider a potential F . The strategy is to decompose the curves $\gamma = \gamma[F]$ and $\sigma = \sigma(F)$ into finitely many small parts and large jumps, depending on δ , each of which have some desirable properties. With this purpose in mind, consider the set of locations of the large jumps, $\mathbb{J} = \{n \in \mathbb{Z} : |F_n| > \frac{\delta}{2}\}$. The following bound for its cardinality, denoted $J = \#\mathbb{J}$, follows from Chebyshev's inequality:

$$J < \left(\frac{2\|F\|_{\ell^r}}{\delta} \right)^r. \quad (2.27)$$

Let $n_1 < n_2 < \dots < n_J$ denote the elements of the set \mathbb{J} . For each $1 \leq j \leq J$, set $\sigma_j^\ell = \sigma|_{[n_{j-1}, n_j]}$ and $\gamma_j^\ell = \gamma|_{[n_{j-1}, n_j]}$. These induce natural decompositions

$$\gamma = \gamma_1^s \oplus \gamma_1^\ell \oplus \gamma_2^s \oplus \gamma_2^\ell \oplus \dots \oplus \gamma_J^s \oplus \gamma_J^\ell \oplus \gamma_{J+1}^s, \quad (2.28)$$

$$\sigma = \sigma_1^s \oplus \sigma_1^\ell \oplus \sigma_2^s \oplus \sigma_2^\ell \oplus \dots \oplus \sigma_J^s \oplus \sigma_J^\ell \oplus \sigma_{J+1}^s. \quad (2.29)$$

By the triangle inequality,

$$\mathcal{V}_r(\gamma) \leq \sum_{j=1}^J \mathcal{V}_r(\gamma_j^\ell) + \sum_{j=1}^{J+1} \mathcal{V}_r(\gamma_j^s). \quad (2.30)$$

Using Hölder's inequality, estimate (2.27), and the fact that $\mathcal{V}_r(\gamma_j^\ell) = \frac{1}{2} \log\left(\frac{1+|F_{n_j}|}{1-|F_{n_j}|}\right)$, we may estimate the contributions coming from the large jumps as follows:

$$\sum_{j=1}^J \mathcal{V}_r(\gamma_j^\ell) \leq J^{\frac{1}{r}} \left(\sum_{j=1}^J \mathcal{V}_r^r(\gamma_j^\ell) \right)^{\frac{1}{r}} \lesssim_r \|F\|_{\ell^r}^{r-1} \left(\sum_{j=1}^J \mathcal{V}_r^r(\gamma_j^\ell) \right)^{\frac{1}{r}} = \frac{\|F\|_{\ell^r}^{r-1}}{2} \left(\sum_{n \in \mathbb{J}} \log\left(\frac{1+|F_n|}{1-|F_n|}\right)^r \right)^{\frac{1}{r}}. \quad (2.31)$$

To estimate the contributions coming from the small jumps, consider the curve $\tilde{\sigma} = \sigma(F\mathbf{1}_{\mathbb{Z} \setminus \mathbb{J}})$ associated to the potential F with the large jumps removed. We claim that the curve $\tilde{\sigma}$ can be subdivided into $L = O(\delta^{-r} \mathcal{V}_r^r(\sigma))$ sub-curves $\{\tilde{\sigma}_j\}_{j=1}^L$, in such a way that $\mathcal{V}_r(\tilde{\sigma}_j) \leq \delta$, for every $1 \leq j \leq L$. More precisely, each curve σ_j^s admits a further decomposition

$$\sigma_j^s = \bigoplus_{k=1}^{L_j} \sigma_{j,k}^s$$

with each $\mathcal{V}_r(\sigma_{j,k}^s) \leq \delta$, where $L := \sum_{j=1}^{J+1} L_j = O(\delta^{-r} \mathcal{V}_r^r(\sigma))$. This can be seen via iterating a decomposition procedure similar to the one used in the proof of [14, Lemma C.2]. The

details are left to the reader. By Step 2, the r -variation of the corresponding curves $\gamma_{j,k}^s$ satisfies

$$\mathcal{V}_r(\gamma_{j,k}^s) = \mathcal{V}_r(\sigma_{j,k}^s) + O(\mathcal{V}_r^2(\sigma_{j,k}^s)).$$

Applying the triangle inequality and Hölder's inequality as before, we estimate

$$\begin{aligned} \sum_{j=1}^{J+1} \mathcal{V}_r(\gamma_j^s) &\leq \sum_{j=1}^{J+1} \sum_{k=1}^{L_j} \mathcal{V}_r(\gamma_{j,k}^s) = \sum_{j=1}^{J+1} \sum_{k=1}^{L_j} (\mathcal{V}_r(\sigma_{j,k}^s) + O(\mathcal{V}_r^2(\sigma_{j,k}^s))) \\ &\leq L^{\frac{1}{r'}} \left(\sum_{j=1}^{J+1} \sum_{k=1}^{L_j} \mathcal{V}_r^r(\sigma_{j,k}^s) \right)^{\frac{1}{r}} + CL^{\frac{1}{r'}} \left(\sum_{j=1}^{J+1} \sum_{k=1}^{L_j} \mathcal{V}_r^{2r}(\sigma_{j,k}^s) \right)^{\frac{1}{r}}. \end{aligned}$$

The reverse triangle inequality and the bound $\mathcal{V}_r(\sigma_{j,k}^s) \leq \delta$ then imply

$$\sum_{j=1}^{J+1} \mathcal{V}_r(\gamma_j^s) \leq L^{\frac{1}{r'}} (1 + C\delta) \mathcal{V}_r(\sigma).$$

Recalling that $L = O(\delta^{-r} \mathcal{V}_r^r(\sigma))$, and that C, δ depend only on r , we have that

$$\sum_{j=1}^{J+1} \mathcal{V}_r(\gamma_j^s) \lesssim_r \mathcal{V}_r^r(\sigma). \quad (2.32)$$

Since $1 < r < 2$, we may then take an appropriate geometric mean to verify that

$$\sum_{j=1}^{J+1} \mathcal{V}_r(\gamma_j^s) \leq \mathcal{V}_r(\sigma) + C_r(\mathcal{V}_r^2(\sigma) \wedge \mathcal{V}_r^r(\sigma)). \quad (2.33)$$

Inequality (2.3) follows from estimates (2.30), (2.31) and (2.33), and this finishes the proof of the proposition. \square

Proposition 2.1 controls the r -variation of the curve γ in terms of that of σ , plus an extra term that accounts for the large jumps. It is reasonable to ask whether, in the converse direction, the r -variation of σ can be controlled by that of γ . An analysis of the previous proof reveals this to be the case, without the need for an extra term as before. We formulate this observation as our next result, which is going to play a role in extending the range of exponents for which Theorem 1.1 holds.

Corollary 2.2. *Let $1 \leq r < 2$. Then there exists a constant $C_r < \infty$ such that*

$$\mathcal{V}_r(\sigma)(z) \leq \mathcal{V}_r(\gamma)(z) + C_r(\mathcal{V}_r^2(\gamma)(z) \wedge \mathcal{V}_r^r(\gamma)(z)), \quad (2.34)$$

for every potential F and every $z \in \mathbb{T}$.

Proof. As observed in the course of the proof of [14, Lemma C.3] via a continuity argument, the fact that estimate (2.26) holds provided $\mathcal{V}_r(\sigma)(z) \leq \delta$, for sufficiently small δ , implies

$$\mathcal{V}_r(\sigma)(z) = \mathcal{V}_r(\gamma)(z) + O_r(\mathcal{V}_r^2(\gamma)(z)),$$

provided $\mathcal{V}_r(\gamma)(z) \leq \delta$. One is thus left with establishing the analogue of Step 3 of the proof of Proposition 2.1. The estimate of the contributions from the small jumps is analogous

and yields

$$\sum_{j=1}^{J+1} \mathcal{V}_r(\sigma_j^s) \lesssim_r \mathcal{V}_r^r(\gamma). \quad (2.35)$$

Since $\|F\|_{\ell^r} \leq \mathcal{V}_r(\gamma)$, we can estimate the contributions from the large jumps as follows:

$$\sum_{j=1}^J \mathcal{V}_r(\sigma_j^\ell) \lesssim_r \|F\|_{\ell^r}^{r-1} \left(\sum_{j=1}^J \mathcal{V}_r^r(\sigma_j^\ell) \right)^{\frac{1}{r}} = \|F\|_{\ell^r}^{r-1} \left(\sum_{n \in \mathbb{J}} |F_n|^r \right)^{\frac{1}{r}} \leq \|F\|_{\ell^r}^r \leq \mathcal{V}_r^r(\gamma). \quad (2.36)$$

The upper bounds given by (2.35) and (2.36) coincide, and the argument can be finished as before to yield (2.34). \square

3. A DISCRETE VARIATIONAL MENSHOV–PALEY–ZYGmund THEOREM

This section is devoted to controlling the variation of the linearized curve $\sigma(F)$ in terms of the potential F . The following result should be compared to the classical Menshov–Paley–Zygmund theorem on the torus.

Proposition 3.1. *Let $1 \leq p < 2$ and $r > p$. Then there exists a constant $C_{p,r} < \infty$ such that*

$$\|\mathcal{V}_r(\sigma(F))\|_{L^{p'}(\mathbb{T})} \leq C_{p,r} \|F\|_{\ell^p(\mathbb{Z})}, \quad (3.1)$$

for every $F \in \ell^p(\mathbb{Z}, \mathcal{D})$.

The condition $r > p$ is sharp, except at the endpoint $p = 1$. In this case, if $p = r = 1$, then inequality (3.1) turns into an equality with $C_{1,1} = 1$. In the general case, (3.1) can only hold if $r \geq p$, as can easily be seen by considering truncated potentials given by $F_n = \frac{1}{2}$ if $|n| \leq M$, and $F_n = 0$ otherwise. These of course coincide with (a multiple of) the Fourier transform of the Dirichlet kernel, defined for $M \geq 1$ to be

$$D_M(x) = \sum_{n=-M}^M e^{inx} = \frac{\sin((M + \frac{1}{2})x)}{\sin \frac{x}{2}}.$$

Moreover, if $p > 1$, then $r > p$ is a necessary condition for (3.1) to hold. This follows from an adaptation of the argument in [14, §2] by testing (3.1) against the Fourier transform of the de la Vallée-Poussin kernel. If $K_M = (M + 1)^{-1} \sum_{n=0}^M D_n$ denotes the Fejér kernel, then set $V_M = 2K_{2M+1} - K_M$, and define the potential $F^M \in \ell^p$ via

$$2F_n^M = \frac{1}{2\pi} \int_0^{2\pi} V_M(x) e^{-inx} dx = \begin{cases} 1, & \text{if } |n| \leq M + 1, \\ \frac{2M+2-|n|}{M+1}, & \text{if } M + 1 \leq |n| \leq 2M + 2, \\ 0, & \text{if } |n| \geq 2M + 2. \end{cases}$$

Since $\|F^M\|_{\ell^p} = O_p(M^{\frac{1}{p}})$, a computation analogous to the one in [14, §2] shows that, for sufficiently large M ,

$$\frac{\|\mathcal{V}_p(\sigma(F^M))\|_{L^{p'}}}{\|F^M\|_{\ell^p}} \geq c_p (\log M)^{\frac{1}{p'}},$$

hence the claimed necessity. For small values of $|F_n|$, the nonlinear Fourier transform is well approximated by the linear Fourier transform, as can be seen by linearizing in F . It follows from this general principle and the discussion in the last paragraph that the

condition $r > p$ is necessary for Theorem 1.1 to hold, at least in the interesting case when $p > 1$.

It seems possible to adapt the high-powered methods from [14] to establish the case $p = 2$ of Proposition 3.1, at which point the result would follow from interpolation with a trivial estimate at $p = 1$. However, the purpose of this section is to provide an elementary proof when $p < 2$ based on embedding $\ell^p(\mathbb{Z})$ into $L^p(\mathbb{R})$.

Proof of Proposition 3.1. For $x, y \in \mathbb{R}$, consider the kernel

$$K(x, y) = \begin{cases} e^{ix[y]} & \text{if } x \in [0, 2\pi], \\ 0 & \text{otherwise,} \end{cases}$$

where $[y]$ denotes the largest integer smaller than y . The associated integral operator is

$$Tf(x) = \int_{\mathbb{R}} K(x, y)f(y) dy.$$

Since $1 \leq p < 2$, the operator T is bounded from $L^p(\mathbb{R})$ to $L^{p'}(\mathbb{R})$, with operator norm satisfying $\|T\|_{p,p'} \leq (2\pi)^{\frac{1}{p'}}$. To verify this, consider the averages $\langle f \rangle_n = \int_n^{n+1} f(y) dy$. Identifying $z = e^{ix}$, we have that

$$\frac{1}{2\pi} \int_{\mathbb{R}} |Tf(x)|^{p'} dx = \int_{\mathbb{T}} \left| \sum_{n \in \mathbb{Z}} \langle f \rangle_n z^n \right|^{p'} = \|\{\langle f \rangle_n\}^\sim\|_{L_z^{p'}(\mathbb{T})}^{p'}.$$

By the (dual) Hausdorff–Young inequality on \mathbb{T} and Hölder's inequality,

$$\frac{1}{2\pi} \int_{\mathbb{R}} |Tf(x)|^{p'} dx \leq \|\{\langle f \rangle_n\}\|_{\ell^p(\mathbb{Z})}^{p'} \leq \|f\|_{L^p(\mathbb{R})}^{p'},$$

and the claimed boundedness of T follows. We proceed to define the truncated operators

$$T_{\leq} f(x, N) = \int_{-\infty}^N K(x, y)f(y) dy.$$

Given a sequence $F \in \ell^p(\mathbb{Z}, \mathcal{D})$, we construct a companion function $f \in L^p(\mathbb{R}, \mathcal{D})$ via

$$f = \sum_{n \in \mathbb{Z}} F_n \mathbf{1}_{[n, n+1)}.$$

For every $0 \leq x \leq 2\pi$ and $N \in \mathbb{Z}$, we have that

$$T_{\leq} f(x, N) = \sum_{n < N} F_n e^{inx}.$$

Writing $z = e^{ix}$ as before, it follows that

$$\|\mathcal{V}_r(\sigma(F))(z)\|_{L_z^{p'}(\mathbb{T})} = \|\mathcal{V}_r(T_{\leq} f)(x)\|_{L_x^{p'}([0, 2\pi])}. \quad (3.2)$$

Since $1 \leq p < \min\{p', r\}$, we are in a position to apply the following result, which is an immediate consequence of [14, Lemma B.1].

Lemma 3.2. *Under the same assumptions on the exponents p, r , there exists a constant $C_{p,r} < \infty$ such that*

$$\|\mathcal{V}_r(T_{\leq} g)\|_{L^{p'}(\mathbb{R})} \leq C_{p,r} \|T\|_{p,p'} \|g\|_{L^p(\mathbb{R})}, \quad (3.3)$$

for every $g \in L^p(\mathbb{R})$.

Since $\|T\|_{p,p'} \leq (2\pi)^{\frac{1}{p'}}$ and $\|f\|_{L^p(\mathbb{R})} = \|F\|_{\ell^p(\mathbb{Z})}$, estimates (3.2) and (3.3) with $g = f$ imply inequality (3.1), concluding the proof of the proposition. \square

Lemma 3.2 is proved in [14] via a bootstrap argument which uses the linear Christ–Kiselev lemma as a black box. We now present an alternative approach which follows the original argument used to establish [2, Theorem 1.1].

Alternative proof of Lemma 3.2. Without loss of generality, we may assume that $\|g\|_{L^p(\mathbb{R})} = 1$ and $r < p'$. For $t \in \mathbb{R}$, define the function

$$\varphi(t) = \int_{-\infty}^t |g(x)|^p dx.$$

Since φ is continuous, non-decreasing, and runs from 0 to 1, we may define points $\{t_k^n\}$ indexed by $n \in \mathbb{N}$ and $1 \leq k < 2^n$ via

$$\varphi(t_k^n) = \frac{k}{2^n},$$

where t_k^n is the smallest such solution if φ has a flat piece. In this range of n, k , define the intervals $I_k^n = (t_{k-1}^n, t_k^n]$, and further set $I_{2^n}^n = \mathbb{R} \setminus \bigcup_{1 \leq k < 2^n} I_k^n$. One easily checks, for every n, k , that

$$\int_{I_k^n} |g(x)|^p dx = 2^{-n}. \quad (3.4)$$

Given $t \in \mathbb{R}$, write $\varphi(t) = \sum_{n=1}^{\infty} j_n 2^{-n}$ with $j_n \in \{0, 1\}$ as a binary expansion. If more than one such expansion exists, choose one. Define $k_n(t) = \sum_{\ell=1}^n j_\ell 2^{n-\ell}$. By continuity of the function φ , we have that

$$g \mathbf{1}_{(-\infty, t]} = \sum_{\{n: j_n=1\}} g \mathbf{1}_{I_{k_n(t)}^n},$$

as functions in L^p . Proceed similarly for intervals of the form $[s, \infty)$. It follows that, for any reals $N_j < N_{j+1}$,

$$|T_{\leq} g(x, N_{j+1}) - T_{\leq} g(x, N_j)| = |T(g \mathbf{1}_{[N_j, N_{j+1}]})(x)| \leq 2 \sum_{n=1}^{\infty} \sup_k |T(g \mathbf{1}_{I_k^n})(x)|,$$

where, for each n , the supremum in the last summand is taken over all $1 \leq k \leq 2^n$ such that $I_k^n \subset [N_j, N_{j+1}]$. As a consequence,

$$\begin{aligned} \left(\sum_{j=0}^{K-1} |T_{\leq} g(x, N_{j+1}) - T_{\leq} g(x, N_j)|^r \right)^{\frac{1}{r}} &\leq 2 \left(\sum_{j=0}^{K-1} \left(\sum_{n=1}^{\infty} \sup_k |T(g \mathbf{1}_{I_k^n})(x)| \right)^r \right)^{\frac{1}{r}} \\ &\leq 2 \sum_{n=1}^{\infty} \left(\sum_{j=0}^{K-1} \sup_k |T(g \mathbf{1}_{I_k^n})(x)|^r \right)^{\frac{1}{r}} \leq 2 \sum_{n=1}^{\infty} \left(\sum_{k=1}^{2^n} |T(g \mathbf{1}_{I_k^n})(x)|^r \right)^{\frac{1}{r}}, \end{aligned}$$

where the last two inequalities are a consequence of Minkowski's inequality and the fact that, for each n , $\{I_k^n : 1 \leq k \leq 2^n\}$ is a partition of \mathbb{R} into disjoint subintervals. Hence

$$\begin{aligned} \|\mathcal{V}_r(T \leq g)(x)\|_{L_x^{p'}(\mathbb{R})} &\leq 2 \left(\int_{\mathbb{R}} \left| \sum_{n=1}^{\infty} \left(\sum_{k=1}^{2^n} |T(g \mathbf{1}_{I_k^n})(x)|^r \right)^{\frac{1}{r}} \right|^{p'} dx \right)^{\frac{1}{p'}} \\ &\leq 2 \sum_{n=1}^{\infty} \left(\int_{\mathbb{R}} \left(\sum_{k=1}^{2^n} |T(g \mathbf{1}_{I_k^n})(x)|^r \right)^{\frac{p'}{r}} dx \right)^{\frac{1}{p'}} \leq 2 \sum_{n=1}^{\infty} \left(\sum_{k=1}^{2^n} \left(\int_{\mathbb{R}} |T(g \mathbf{1}_{I_k^n})(x)|^{p'} dx \right)^{\frac{r}{p'}} \right)^{\frac{1}{r}}, \end{aligned}$$

where Minkowski's integral inequality was used to interchange the ℓ^r and the $L^{p'}$ norms (recall that $r < p'$). Since the operator T is bounded from L^p to $L^{p'}$, we appeal to (3.4) to further estimate

$$\|\mathcal{V}_r(T \leq g)(x)\|_{L_x^{p'}(\mathbb{R})} \leq 2 \|T\|_{p,p'} \sum_{n=1}^{\infty} \left(\sum_{k=1}^{2^n} 2^{-\frac{nr}{p}} \right)^{\frac{1}{r}}.$$

Since $r > p$, the geometric series converges, concluding the proof of the lemma with $C_{p,r} = 2(2^{\frac{1}{p}-\frac{1}{r}} - 1)^{-1}$. \square

4. PROOF OF THE MAIN THEOREM

Armed with Propositions 2.1 and 3.1, it is now an easy matter to finish the proof of the main theorem.

Proof of Theorem 1.1. Let $1 \leq p < 2$ and $r > p$. We deal with the case $r < 2$ first. In this case, given a potential F , we start by considering the set

$$\mathbb{S} = \{z \in \mathbb{T} : \mathcal{V}_r(\gamma[F])(z) \leq 1\},$$

defined in (1.9). For any point $z \in \mathbb{S}$, we have that $\mathcal{V}_r(\gamma[F])(z) \leq 1$, and therefore $\mathcal{V}_r(\sigma[F])(z) \lesssim_r 1$ in view of Corollary 2.2. As observed before, this implies $\|F\|_{\ell^r} \lesssim_r 1$. From Proposition 2.1, we obtain

$$\mathcal{V}_r(\gamma[F])(z) \lesssim_r \mathcal{V}_r(\sigma(F))(z) + \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^r}.$$

From Proposition 3.1, the elementary inequality $x \leq \frac{1}{2} \log \left(\frac{1+x}{1-x} \right)$, and the inclusion $\ell^p \subset \ell^r$, it follows that

$$\|\mathcal{V}_r(\gamma[F])\|_{L^{p'}(\mathbb{S})} \lesssim_{p,r} \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^p(\mathbb{Z})}. \quad (4.1)$$

We now consider the complementary set $\mathbb{T} \setminus \mathbb{S}$. Given $z \in \mathbb{T} \setminus \mathbb{S}$, estimate (2.3) implies $\mathcal{V}_r(\sigma[F])(z) \gtrsim_r 1$, and therefore

$$\mathcal{V}_r(\gamma[F])(z) \lesssim_r \mathcal{V}_r^r(\sigma(F))(z) + \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^r}^r.$$

This leads to

$$\|\mathcal{V}_r(\gamma[F])\|_{L^{\frac{p'}{r}}(\mathbb{T} \setminus \mathbb{S})}^{\frac{1}{r}} \lesssim_r \|\mathcal{V}_r^r(\sigma(F))\|_{L^{\frac{p'}{r}}(\mathbb{T})}^{\frac{1}{r}} + \left\| \log \left(\frac{1 + |F_n|}{1 - |F_n|} \right) \right\|_{\ell^r(\mathbb{Z})}.$$

Another application of Proposition 3.1 shows that

$$\|\mathcal{V}_r^r(\sigma(F))\|_{L^{\frac{p'}{r}}(\mathbb{T})}^{\frac{1}{r}} = \|\mathcal{V}_r(\sigma(F))\|_{L^{p'}(\mathbb{T})} \lesssim_{p,r} \|F\|_{\ell^p(\mathbb{Z})},$$

and so, reasoning as before,

$$\|\mathcal{V}_r(\gamma[F])\|_{L^{\frac{p'}{r}}(\mathbb{T}\backslash\mathbb{S})}^{\frac{1}{r}} \lesssim_{p,r} \left\| \log \left(\frac{1+|F_n|}{1-|F_n|} \right) \right\|_{\ell^p(\mathbb{Z})}. \quad (4.2)$$

Estimates (4.1) and (4.2) yield the desired inequality (1.8), as long as $r < 2$.

In order to extend it to values $r \geq 2$, we proceed as follows. Fix $1 \leq p < 2$ and $r \geq 2$. Set $s = \frac{p+2}{2}$ and $\mathbb{S} = \{z \in \mathbb{T} : \mathcal{V}_s(\gamma[F])(z) \leq 1\}$. Since $s < 2$, the preceding discussion implies

$$\|\mathcal{V}_s(\gamma[F])\|_{L^{p'}(\mathbb{S})} + \|\mathcal{V}_s^{\frac{1}{s}}(\gamma[F])\|_{L^{p'}(\mathbb{T}\backslash\mathbb{S})} \lesssim_p \left\| \log \left(\frac{1+|F_n|}{1-|F_n|} \right) \right\|_{\ell^p(\mathbb{Z})}.$$

If $z \in \mathbb{S}$, then it suffices to recall that the r -variation is decreasing in the exponent r :

$$\mathcal{V}_r(\gamma[F])(z) \leq \mathcal{V}_s(\gamma[F])(z)$$

If $z \in \mathbb{T}\backslash\mathbb{S}$, then we may estimate:

$$\mathcal{V}_r^{\frac{1}{r}}(\gamma[F])(z) \leq \mathcal{V}_s^{\frac{1}{r}}(\gamma[F])(z) \leq \mathcal{V}_s^{\frac{1}{s}}(\gamma[F])(z),$$

where the first inequality follows from the fact that the r -variation is decreasing in the exponent r , and the second inequality follows from the fact that $\mathcal{V}_s(\gamma[F])(z) > 1$.

As a consequence, inequality (1.8) holds for arbitrary potentials, in the full range $1 \leq p < 2$ and $r > p$. Since the implicit constant depend only on the exponents p and r , the result for general ℓ^p sequences follows. The proof of the theorem is complete. \square

ACKNOWLEDGMENTS

The author would like to thank Christoph Thiele for suggesting the problem and for several stimulating discussions. He is also indebted to Vjeko Kovač, João Pedro Ramos and Gena Uraltsev for various comments and suggestions, and to Jim Wright for an enlightening conversation and a copy of an unpublished joint work with Sandy Davie.

REFERENCES

- [1] Michael Christ and Alexander Kiselev, *Absolutely continuous spectrum for one-dimensional Schrödinger operators with slowly decaying potentials: some optimal results*. J. Amer. Math. Soc. **11** (1998), no. 4, 771–797.
- [2] Michael Christ and Alexander Kiselev, *Maximal functions associated to filtrations*. J. Funct. Anal. **179** (2001), no. 2, 409–425.
- [3] Yen Do, Camil Muscalu and Christoph Thiele, *Variational estimates for paraproducts*. Rev. Mat. Iberoam. **28** (2012), no. 3, 857–878.
- [4] Yen Do, Camil Muscalu and Christoph Thiele, *Variational estimates for the bilinear iterated Fourier integral*. J. Funct. Anal. **272** (2017), no. 5, 2176–2233.
- [5] Roger L. Jones, Andreas Seeger and James Wright, *Strong variational and jump inequalities in harmonic analysis*. Trans. Amer. Math. Soc. **360** (2008), no. 12, 6711–6742.
- [6] Vjekoslav Kovač, *Uniform constants in Hausdorff–Young inequalities for the Cantor group model of the scattering transform*. Proc. Amer. Math. Soc. **140** (2012), no. 3, 915–926.

- [7] Vjekoslav Kovač, Diogo Oliveira e Silva and Jelena Rupčić, *A sharp nonlinear Hausdorff–Young inequality for small potentials*. Preprint (2017), available at arXiv:1703.05557.
- [8] Allison Lewko and Mark Lewko, *Estimates for the square variation of partial sums of Fourier series and their rearrangements*. J. Funct. Anal. **262** (2012), no. 6, 2561–2607.
- [9] Terry Lyons, *Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young*. Math. Res. Lett. **1** (1994), no. 4, 451–464.
- [10] Terry Lyons, *Differential equations driven by rough signals*. Rev. Mat. Iberoam. **14** (1998), no. 2, 215–310.
- [11] Terry Lyons and Danyu Yang, *The partial sum process of orthogonal expansions as geometric rough process with Fourier series as an example—an improvement of Menshov–Rademacher theorem*. J. Funct. Anal. **265** (2013), no. 12, 3067–3103.
- [12] Camil Muscalu, Terence Tao and Christoph Thiele, *A counterexample to a multilinear endpoint question of Christ and Kiselev*. Math. Res. Lett. **10** (2003), no. 2-3, 237–246.
- [13] Camil Muscalu, Terence Tao and Christoph Thiele, *A Carleson theorem for a Cantor group model of the scattering transform*. Nonlinearity **16** (2003), no. 1, 219–246.
- [14] Richard Oberlin, Andreas Seeger, Terence Tao, Christoph Thiele and James Wright, *A variation norm Carleson theorem*. J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 421–464.
- [15] Barry Simon, *Orthogonal polynomials on the unit circle. Part 2. Spectral theory*. American Mathematical Society Colloquium Publications, 54, Part 2. American Mathematical Society, Providence, RI, 2005.
- [16] Terence Tao and Christoph Thiele, *Nonlinear Fourier Analysis*. IAS/Park City Graduate Summer School, unpublished lecture notes (2003), available at arXiv:1201.5129.
- [17] Norbert Wiener, *The quadratic variation of a function and its Fourier coefficients*. MIT Journal of Math. and Physics **3** (1924), 72–94.

DIOGO OLIVEIRA E SILVA, HAUSDORFF CENTER FOR MATHEMATICS, 53115 BONN, GERMANY
E-mail address: dosilva@math.uni-bonn.de